Looking Backward and Forward on Hindsight Bias

Daniel M. Bernstein, Andre Aßfalg, Ragav Kumar, and Rakefet Ackerman

Abstract

The same event that appeared unpredictable in foresight can be judged as predictable in hindsight. Hindsight bias clouds judgments in all areas of life, including legal decisions, medical diagnoses, consumer satisfaction, sporting events, and election outcomes. We discuss three theoretical constructs related to hindsight bias: memory, reconstruction bias, and motivation. Attempts to recall foresight knowledge fail because newly acquired knowledge affects memory either directly or indirectly by biasing attempts to reconstruct foresight knowledge. On a metacognitive level, overconfidence and surprise contribute to hindsight bias. Overconfidence in knowledge increases hindsight bias whereas a well-calibrated confidence reduces hindsight bias. Motivational factors also contribute to hindsight bias by making positive and negative outcomes appear more or less likely, depending on a variety of factors. We review hindsight bias theories and discuss three exciting directions for future research.

Key words: hindsight bias, metacognition, memory, motivation, theory of mind, learning

On March 31, 2009, six Italian scientists and a former government official met in the ancient city of L’Aquila to discuss the possibility of a major earthquake in the region. Despite recent tremors recorded nearby, the group concluded that it was impossible to predict a major earthquake. Six days later, a 6.3 magnitude quake devastated the ancient city, killing 300 people. On October 22, 2012, an Italian court sentenced the unlucky group to six years in jail for manslaughter for the group’s failure to warn the public of the pending quake. The court ordered the group to pay 7.8 million euros in damages (The Telegraph, 2012).

Judgments under uncertainty, like predicting earthquakes, are hard at the best of times (Kahneman, 2011; Tversky & Kahneman, 1974). Without outcome knowledge, the scientists and former government official made an educated guess about the future. With outcome knowledge, the Italian court likely made an overeducated guess, claiming that the group “should have known” more than they truly knew. Hindsight bias makes uncertain events seem predictable and inevitable (Fischhoff, 1975).
Baruch Fischhoff was the first to study hindsight bias experimentally. As a graduate student caught up in Tversky and Kahneman’s (1973, 1974) scientific juggernaut of heuristics and biases, Fischhoff and his fellow graduate students read an article by Paul Meehl on “clinicians’ exaggerated feeling of having known all along how cases were going to turn out” (Fischhoff, 2007; p. 10). Fischhoff immediately linked Meehl’s observation to politics, in particular how, in hindsight, outcomes to political events make those events seem far more inevitable than they were in foresight. To test this hypothesis, Fischhoff and Beyth (1975) asked participants to rate the likelihood of various outcomes to then–US President Nixon’s pending trip to China and the Soviet Union. For example, participants rated the likelihood that Nixon would meet Chairman Mao and that Nixon would declare the trip a success. After Nixon completed the trip, Fischhoff and Beyth asked participants to recall their initial predictions. The results were clear: Participants gravitated toward the actual event outcomes. For instance, participants who initially thought it unlikely that Nixon would meet Mao later recalled that they had thought this meeting likely. This study, and another one by Fischhoff (1975), launched the scientific study of hindsight bias.

Hindsight bias has many aliases. A sample of these include creeping determinism (Fischhoff, 1975), the I-knew-it-all-along effect (Wood, 1978), outcome bias (Baron & Hershey, 1988), the curse of knowledge (Camerer, Lowenstein, & Weber, 1989), mental contamination (Wilson & Brekke, 1994), realist bias (Mitchell, Robinson, Isaacs, & Nye, 1996), adult egocentrism (Kelley & Jacoby, 1996), and epistemic egocentrism (Royzman, Cassidy, & Baron, 2003). Although some authors distinguish these variants from hindsight bias, all variants have one thing in common: It is hard to ignore privileged information when trying to reason from a naïve perspective.

Hindsight bias is a common error that occurs in many domains, including legal decisions and medical diagnoses (Arkes, Wortmann, Saville, & Harkness, 1981; Harley, 2007), consumer satisfaction (Zwick, Pieters, & Baumgartner, 1995), sporting events, and election outcomes (Leary, 1981, 1982). In each case, advance knowledge of an outcome causes people to overestimate the outcome’s likelihood (see Arkes, 2013; Hawkins & Hastie, 1990; Roese & Vohs, 2012 for reviews). In applied terms, hindsight bias causes people to overestimate what other people know. This results in miscommunication when a speaker overestimates the clarity of her message or when a listener overestimates her understanding of a message (see Bernstein, Wilson, Pernat, & Meilleur, 2012). In a classroom, such miscommunication results in ineffective teaching and learning. In writing, such miscommunication results in poor delivery of ideas, and poor understanding on the reader’s part. In essence, one should avoid hindsight bias to communicate effectively.

Hindsight bias exists across cultures (Pohl, Bender, & Lachman, 2002; although see Heine & Lehman, 1996; Wu & Keysar, 2007) and has been documented across the lifespan (Bernstein, Erdfelder, Meltzoff, Peria, & Loftus, 2011; see Bayen, Pohl, Erdfelder, & Auer, 2007; Birch & Bernstein, 2007 for reviews). The few developmental studies to date show that children, adults, and the elderly all exhibit hindsight bias, but preschoolers and the elderly are more prone to it (Bayen, Erdfelder, Bearden, & Lozito, 2006; Bernstein et al., 2011; Coolin, Bernstein, Thornton, &
Hindsight bias has been shown using different types of stimuli, from general-knowledge almanac questions to event outcomes, including visual, auditory, and gustatory judgments (Bernstein et al., 2012; Harley, Carlsen, & Loftus, 2004; Lange, Thomas, Dana, & Dawes, 2011; Pohl, Schwarz, Sczesny, & Stahlberg, 2003).

Researchers examine hindsight bias mainly by using one of two experimental designs: the memory design and the hypothetical design. In the memory design, depicted in the left panel of Figure 1, participants make a baseline judgment about a topic, expected to be unknown to the majority of the target population. This judgment may include a probability assessment about the outcome to an event or answer to an almanac question that elicits a numerical answer (length, year, amount, etc.; e.g., Hell, Gigerenzer, Gauggel, Mall, and Müller, 1988). For example, when asked how many keys there are on a standard piano keyboard, a participant might guess, “50.” Upon later learning that it has 88 keys, and when asked what her original guess was, the now-knowledgeable participant may say, “80.” The numerical nature of these questions allows researchers to evaluate the extent to which the participant has gravitated toward the correct answer when attempting to recall her original answer. Such answers are then contrasted with control questions, to which the participant does not learn the correct answer before recalling her original answer. The classic hindsight bias finding is that participants more accurately recall their original answer to the control questions, and if they fail to recall it, they move randomly toward and away from the correct answer.

The hindsight bias literature also makes use of another type of question, which involves assessment of whether a particular event will happen. For example, participants predict the probability of one possible outcome to a sporting event or election. In these studies, participants who know the actual outcome often claim in hindsight, that they had predicted this outcome (Leary, 1981, 1982). The memory design elicits two different measures of hindsight bias: occurrence and magnitude. Hindsight bias occurrence reflects the number of times a person’s recalled original judgments shift toward the correct answers. Hindsight bias magnitude reflects the size of bias or the degree to which incorrectly recalled original judgments gravitate toward the correct answers (Coolin et al., 2014).

In the hypothetical design, depicted in the right panel of Figure 1, a participant learns the correct answer to an almanac question (e.g., a standard piano keyboard has 88 keys) and then estimates what she would have answered had she not known the correct answer (see Campbell & Tesser, 1983). As with the memory design, the hypothetical design also contains control questions to which the participant does not learn the correct answer. Again, knowing the correct answer colors one’s judgment for a naïve self, causing the participant to say that she

---

1 Such tasks are also used for studying flexibility in the regulation of what is called grain size—the specificity with which a person answers questions—in contexts such as memory processes and decision making in social communication and forensic investigations (e.g., Ackerman & Goldsmith, 2008; Goldsmith, Koriat, & Weinberg-Eliezer, 2002; Haran, Moore, & Morewedge, 2010; McKenzie, Liersch, & Yaniv, 2008; Weber & Brewer, 2008).
Looking Backward and Forward on Hindsight Bias, Bernstein et al. (in press)

would have answered, “80 keys” when in actuality, the naïve person’s guesses would be around 50. Sometimes participants estimate for a naïve peer instead of for themselves (e.g., Wood, 1978). Participants might be asked, for example, how many keys a naïve peer would estimate are on a standard piano keyboard. Just as with the hypothetical-self judgment, estimates for a naïve other when one knows the correct answer resemble the correct answer more than estimates for a naïve other when one does not know the correct answer.

*Figure 1.* The memory (left panel) and hypothetical (right panel) design in hindsight bias experiments.

To understand the cognitive processes that underlie hindsight bias, researchers have proposed that there are three distinct, hierarchically organized components of hindsight bias (see Blank, Nestler, von Collani, & Fischer, 2008; Roese & Vohs, 2012). See Table 1. Our review in the following sections refers to these components.

Table 1.
Three-component model of hindsight bias (adapted from Figure 1 in Roese and Vohs, 2011)

1. Memory distortion (“I said it would happen”)
2. Impressions of inevitability (“It had to happen”)
3. Impressions of foreseeability (“I knew it would happen”)
We separate our chapter into two main sections. In Looking Backward, we review several hindsight bias theories (see Table 2). In Looking Forward, we discuss three exciting directions that hindsight bias researchers might pursue in future work.

Looking Backward

Over the years, many authors have suggested explanations for hindsight bias, thereby making it difficult to keep track of all the relevant theories. Instead of discussing each hindsight bias theory in turn, along with empirical evidence for and against, we order our analysis by theoretical constructs supposedly involved in hindsight bias. These constructs are often shared between several theories and include memory, biased reconstruction of the original answer, and motivation (see Table 2).
Table 2.
List of hindsight bias theories related to three theoretical constructs, memory, reconstruction bias, and motivation, discussed in this chapter

<table>
<thead>
<tr>
<th>Memory</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Theory</strong></td>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>Automatic assimilation</td>
<td>The encoding of the correct answer is automatic, unconscious, or</td>
</tr>
<tr>
<td></td>
<td>effortless</td>
</tr>
<tr>
<td>Trace-strength hypothesis</td>
<td>Memory traces for the original and correct answers coexist; their</td>
</tr>
<tr>
<td></td>
<td>relative strength determines the size of hindsight bias</td>
</tr>
<tr>
<td>Recollection bias</td>
<td>Learning the correct answer decreases the chance of recalling the</td>
</tr>
<tr>
<td></td>
<td>original answer</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reconstruction bias</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Theory</strong></td>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>Anchoring and adjustment</td>
<td>Participants anchor their response to the correct answer and</td>
</tr>
<tr>
<td></td>
<td>insufficiently adjust toward the original answer</td>
</tr>
<tr>
<td>Biased memory search</td>
<td>The correct answer guides the memory search toward content related</td>
</tr>
<tr>
<td></td>
<td>to the correct answer</td>
</tr>
<tr>
<td>Metacognitive regulation</td>
<td>Adjustment toward the correct answer depends on confidence in the</td>
</tr>
<tr>
<td></td>
<td>original and recalled answers and on the extent of surprise caused</td>
</tr>
<tr>
<td></td>
<td>by the newly learned information</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Motivation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Theory</strong></td>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>Retroactive pessimism</td>
<td>Manage disappointment by making negative outcomes seem more</td>
</tr>
<tr>
<td></td>
<td>likely and positive outcomes less likely</td>
</tr>
<tr>
<td>Defensive processing</td>
<td>Perceiving the event as less predictable protects one from the</td>
</tr>
<tr>
<td></td>
<td>implication that the outcome was a consequence of one’s own bad</td>
</tr>
<tr>
<td></td>
<td>decisions</td>
</tr>
<tr>
<td>Motivated sense-making</td>
<td>Attempt to make sense of self-relevant outcomes and attribute</td>
</tr>
<tr>
<td></td>
<td>incongruities in expectations to either external reasons (retroactive</td>
</tr>
<tr>
<td></td>
<td>pessimism) or internal reasons (defensive processing)</td>
</tr>
</tbody>
</table>
Memory

In the memory design, participants attempting to recall the original answer after learning the correct answer are biased toward the correct answer. Naturally, explanations of hindsight bias in the memory design postulate that some aspect of memory contributes to the bias.

Fischhoff (1975) argued for an automatic assimilation of the correct answer into memory that is probably effortless and unconscious. Consistent with the idea of automatic assimilation of the correct answer, hindsight bias occurs even when participants are warned about and asked to avoid hindsight bias (Fischhoff, 1977; see also Bernstein et al., 2012; Guilbault, Bryant, Brockway, & Posavac, 2004; Harley et al., 2004; Lilienfeld, Amirati, & Landfield, 2009; Pohl & Hell, 1996). However, later studies have suggested that hindsight bias is less inevitable than Fischhoff (1975) initially argued. For example, discrediting the correct answer eliminates hindsight bias (Erdfeelder & Buchner, 1998; Hasher, Attig, & Alba, 1981). In addition, there is evidence that assimilation of the correct answer into memory could be effortful rather than effortless. Calvillo (2012) found less hindsight bias when participants had to rehearse a four-consonant sequence while receiving the correct answer than without rehearsal. Similarly, Nestler, Blank, and von Collani (2008) had their participants rehearse two- versus eight-digit numbers. Hindsight bias appeared only under low cognitive load conditions (two-digit number) but not under high cognitive load conditions (eight-digit number). It should be noted, however, that participants in these studies were under cognitive load while they received the correct answer and while they tried to recall their original answer. Thus, it is not clear whether cognitive load affected the assimilation of the correct answer, the attempt to recall the original answer, or both.

Hell et al. (1988) proposed a different role of memory in hindsight bias. According to their trace strength hypothesis, the memory traces of the original answer and the correct answer coexist in memory and the amount of hindsight bias depends on the relative memory trace strengths for the original answer and the correct answer. The weaker the memory trace for the original answer relative to the memory trace for the correct answer, the larger the hindsight bias. Supporting this hypothesis, Wood (1978) found that hindsight bias increases when participants receive the correct answer thrice compared to just once (see also Harley et al., 2004). Further, an attempt to recall the original answer one week—as compared to immediately—after receiving the correct answer leads to a smaller hindsight bias (Hell et al., 1988; see also Erdfeelder & Buchner, 1998). This is consistent with the trace-strength hypothesis: If longer retention intervals increase the likelihood of forgetting the correct answer, the trace strength of the original answer increases relative to the trace strength of the correct answer. The result is a weaker hindsight bias for long compared to short retention intervals.

Erdfeelder and Buchner (1998) suggested another role of memory in hindsight bias: According to them, learning the correct answer reduces the chance of retrieving the original answer from long-term memory. This recollection bias might either occur because the correct answer overwrites the memory trace of the original answer or because the correct answer leaves the
memory trace of the original answer intact but its accessibility is weaker than that of the later presented correct answer. The term, “accessibility” in this context pertains to a particular heuristic cue for feeling of knowing (FOK), which reflects the amount of information that one retrieves when encountering a question (see Koriat, 1993, 1995). Indeed, there is evidence that hindsight bias in the memory design is due, in part, to a recollection bias. Erdfelder, Brandt, and Bröder (2007) used a formal model to disentangle recollection bias and reconstruction bias based on the participant’s knowledge and found that recollection bias contributes to hindsight bias—participants were less likely to recall the original answer when they received the correct answer than when they did not receive the correct answer.

Reconstruction Bias

Most authors agree that failing to recollect the original answer is a necessary condition for hindsight bias in the memory design, whereas the bias itself is the result of the participants’ attempt to reconstruct the original answer when retrieval fails (Erdfelder & Buchner, 1998; Hoffrage, Hertwig, & Gigerenzer, 2000; Schwarz & Stahlberg, 2003; Winman, Juslin, & Björkman, 1998). This reconstruction is very much like a cognitive simulation of the foresight condition, where the correct answer was still unavailable (Winman et al., 1998). According to this idea, in hindsight, people try to reconstruct their foresight knowledge. Common to all these theories, the reconstruction process is biased toward the correct answer. Thus, hindsight bias is not an “all or none” phenomenon, with only the chance for correct recollection as an indication; rather, hindsight bias can be gradual. That is, people can be more or less biased regarding the same item. The size of adjustment from their original answer toward the correct answer is the extent of their hindsight bias.

Fischhoff (1975) suggested several cognitive heuristics as possible causes of the biased reconstruction, based on Tversky and Kahneman’s (1974) work. The correct answer might be highly available in memory, appear representative of the judgment domain, or serve as an anchor from which participants try to adjust their reconstruction toward the original answer. For example, asked about the height of the Eiffel Tower a participant might answer, “100 meters.” After learning the correct answer (324 meters) the participant might anchor her response to the correct answer and then adjust toward the original answer (see Epley & Gilovich, 2006). However, because adjustment often fails to be complete (Tversky & Kahneman, 1974), the participant reconstructs the original answer as somewhere between her original answer and the anchor (e.g., 200 meters).

Pohl, Eisenhauer, and Hardt (2003) developed the selective activation and reconstructive anchoring (SARA) model to explain anchoring and adjustment effects as well as hindsight bias. In this model, the reconstruction of the original answer can be biased because the encoding of the correct answer changes the pattern of associations between memory traces and/or because the correct answer guides the memory search during the reconstruction of the original answer toward material related to the correct answer. Interestingly, the SARA model can explain why
the content of the correct answer moderates hindsight bias, a result with which other hindsight bias theories struggle.

For example, Yopchick and Kim (2012) presented participants in the control condition with several scenarios, such as a fictional battle between the Hutu and the Tutsi (see Figure 2). In the hindsight condition, participants additionally received the “correct” answer (e.g., the Hutu won the battle) and information that was either relevant (e.g., the Hutu had superior troop discipline) or irrelevant (e.g., the Hutu started the battle by marching west) to the scenario. The participants’ task was to indicate the likelihood of both outcomes (the Hutu won vs. the Tutsi won) from the viewpoint of a naïve participant who had no access to the correct answer. Interestingly, only high-relevance information led to hindsight bias—that is, participants in the hindsight condition thought the correct answer appeared more likely from the viewpoint of a naïve participant than did participants in the control condition. However, when the correct answer (e.g., the Tutsi won the battle) was surprising given the high-relevance information (e.g., the Hutu had superior troop discipline), a reverse hindsight bias occurred—that is, participants indicated that naïve participants would consider the incorrect answer as more likely than the correct answer (see also Wasserman, Lempert, & Hastie, 1991). In line with this data pattern, SARA suggests that information relevant to the correct answer is likely to serve as a retrieval cue for information in long-term memory that supports the correct answer, biasing the reconstruction of a previous (memory design) or naïve (hypothetical design) view in the direction of the correct answer. Conversely, irrelevant information is likely to serve as a retrieval cue for other irrelevant information resulting in no bias. Finally, information supporting the incorrect answer is likely to cue the retrieval of information supporting the incorrect answer, thus leading to a bias against the correct answer (i.e., a reversed hindsight bias).
Another reconstruction theory, which could be considered a type of biased memory search, is Reconstruction After Feedback with Take the best (RAFT) (Hoffrage et al., 2000). It pertains to
instances in which the participant is attempting to remember an original answer that cannot be retrieved from memory; the participant attempts to reconstruct the original answer and then automatically updates the knowledge used to reconstruct it with all available information including the correct answer (Hertwig, Fanselow, & Hoffrage, 2003; Hoffrage et al., 2000). Hoffrage et al. (2000) argue that, at least when comparing two alternatives (e.g., “Which city is larger, Hamburg or Heidelberg?”), this is an adaptive behavior meant to keep our knowledge current; however, it has a side-effect—because the knowledge used to reconstruct the original answer has been updated, the reconstructed original answer will be influenced by the correct answer, resulting in hindsight bias. Blank and Nestler (2007) note that the RAFT model was developed with pair comparisons in mind, in which a participant must compare two alternatives and make a judgment.

In contrast, metacognitive hindsight bias theories commonly postulate that the reconstruction process is guided by an assessment of the ability to answer questions correctly. For example, Schwarz and Stahlberg (2003) suggested that when people do not remember their original answer, they reconstruct a plausible answer by using the correct answer as an anchor, as just described (see also Werth, Strack, & Förster, 2002). However, the authors demonstrated that the extent of the anchor’s effect depends on one’s perceived knowledge level. Participants who were told that their original answers were close to the correct answers anchored more to the correct answers and thus showed a greater hindsight bias than participants who were told that their original answers were overestimations or underestimations of the correct answers.

Metacognitive judgments can influence hindsight bias both by confidence in the original answers and by hindsight confidence in the success of recalling the original answer (Werth et al., 2002; Winman et al., 1998). The effect of these judgments on the susceptibility to hindsight bias depends on the extent of overconfidence—the gap between the mean confidence across items and actual success rate. For example, Hoch and Loewenstein (1989) found hindsight bias for medium- and high-question difficulty questions but not for easy questions. They explained that although the participants acknowledge that the question is difficult, after learning the correct answer, they fail to appreciate how overconfident they would have been initially without that knowledge and anchor their answer to the presently known correct answer. According to the accuracy-assessment model (Winman et al., 1998), if participants are well calibrated when providing the original answers, they have a higher chance of recalling their answer correctly and also to be well calibrated in hindsight. Overconfidence in the original answers relates to the classic hindsight bias and the feeling that “I knew it all along,” which in turn relates to a high chance of remembering the correct answer. This was termed hypercorrection of high-confidence errors (Butterfield & Metcalfe, 2001; Sitzman, Rhodes, & Tauber, 2014). One interesting explanation for this finding is that when participants are overconfident regarding an incorrect original answer, they are also familiar with the correct answer (Metcalfe & Finn, 2011). See Figure 3. This advance familiarity, in turn, helps one learn the correct answer.
A question of interest is how people infer their own confidence. In general, people infer their own confidence on the basis of experiential cues that suggest the likelihood that they will succeed in the task (Koriat, 1997). With respect to hindsight bias, people may process the correct answer fluently and fail to appreciate their earlier (or naïve others’) lack of knowledge (Harley et al., 2004; Kelley & Jacoby, 1996; Roese & Vohs, 2012). Fluency variability may result from the level of familiarity that one experiences when learning the correct answer (Hoch & Loewenstein, 1989). If the correct answer is not familiar, it may be surprising, and this eliminates the experience of “knowing it all along” and may even reverse the hindsight bias, resulting in responses even farther from the correct answer than naïve respondents would provide (Müller & Stahlberg, 2007). Another basis for the judgment is the effort involved in accessing the original answer and potential alternative answers (Sanna, Schwarz, & Small, 2002). As in other heuristic-based processes, these cues may be misleading. For example, similar to many other fluency-based judgments, hindsight bias was more frequent when the questions were displayed in colors that were easier rather than harder to read (Werth & Strack, 2003; see also Bernstein & Harley, 2007). It thus seems that hindsight bias is generated by combining content knowledge and experiential cues for guiding a metacognitive regulation of answer formation (see Sanna & Schwarz, 2007 for an integrative model).

Pezzo (2003) developed a sense-making model to explain how surprise (that is, how surprising the participant found the correct answer to a question to be) affects hindsight bias. He distinguished between “initial” and “resultant” surprise. Initial surprise occurs when the answer (or, in the case of events, the outcome) is incongruent with prior expectations. Pezzo argued
that if there is no initial surprise, and the correct answer matches expectations, then there is no room for surprise or hindsight bias; it is likely that the participant was either correct or close to being correct. For items where initial surprise exists, the participant then attempts to make sense of the answer to try and rationalize the outcome. If this fails—that is, if the participant cannot make sense of the correct answer—then the correct answer or outcome has high “resultant” surprise, which produces little or no hindsight bias or even produces a reverse hindsight bias (i.e., underestimating their previous ability to respond to the question or predict the outcome correctly: “I never would have known that!”). Conversely, when the sense-making process succeeds, the participant experiences full hindsight bias; in this case, the participant determines that the correct answer was something that could have been reasoned out, and as such believes it to be more likely that she correctly answered the question (or would have correctly answered the question, in hypothetical designs).

Causal model theory (CMT) adapts and extends sense-making theory in order to explain hindsight bias for probabilities of event outcomes (Blank & Nestler, 2007). According to CMT, when a participant learns the outcome to an event that is incongruent with her prior expectations—thereby generating initial surprise—she attempts to make sense of it by elaborating the antecedents of the event and establishing causal relations (which may or may not have any objective validity). This is done by selectively activating supporting evidence and suppressing or deprioritizing conflicting evidence. These causal relations then lead to an increase in the perceived likelihood of the event, perhaps even giving a sense of inevitability; this is what is demonstrated in what Fischhoff (1975) called “creeping determinism.” In terms of Pezzo’s (2003, 2011) sense-making approach, these causal relations are how the participant makes sense of the answer, thus reducing resultant surprise and thereby enhancing hindsight bias.

Within the three-component model of hindsight bias (see Table 1), models like SARA and RAFT deal mainly with memory distortions, while CMT covers impressions of inevitability (Blank & Nestler, 2007). This relates to the other ways in which Blank and Nestler claim that CMT differs from SARA and RAFT; as noted earlier, Nestler et al. (2008) demonstrated that the sense-making process is effortful. Because CMT relies on this process, it follows that CMT is also an effortful, conscious process, while both SARA and RAFT are based on automatic processes during memory encoding and retrieval.

Although there is some consensus about the biased reconstruction of memory traces, the effect of surprise on hindsight bias is particularly difficult to reconcile with many cognitive theories of hindsight bias. Adapting a procedure developed by Roese and colleagues (Roese, Fessel, Summerville, Kruger, & Dilch, 2006), Calvillo and Gomes (2011) had participants watch traffic-related animations, some of which depicted a car crash. The participants’ task was to stop the animation as soon as they were certain that a crash would occur. In the hindsight condition, participants first watched the animation to learn the outcome and then, in a second viewing, rated when a naïve peer would stop the animation (hypothetical design). Car crashes that were rated by an independent group of participants as unsurprising elicited no hindsight bias, while...
medium surprise levels elicited hindsight bias—that is, participants thought a naïve peer would recognize an imminent crash sooner than they themselves did. Conversely, when car crashes were rated as highly surprising, a reverse hindsight bias occurred—participants indicated that a naïve peer would recognize the car crash later than they themselves did.

In summary, hindsight bias is more likely to occur when participants perceive their original answers as generally correct. However, uncertainty should be involved for reconstruction to occur and potentially to generate a bias by failing to ignore the correct answer. One form this uncertainty can take is surprise, wherein the mismatch between expectation and outcome creates potential for hindsight bias, depending on the participants’ ability to make sense of the discrepancy. This role of uncertainty suggests that metacognitive judgments are involved in hindsight bias.

Motivational Theories

Experiments in which hindsight bias is examined by using self-relevant scenarios often reveal a protective function of hindsight bias. This takes place, for example, when participants are asked to imagine a scenario where they may end up missing a flight or have a chance to lose money. Retroactive pessimism examines the use of hindsight bias to manage disappointment (Tykocinski & Steinberg, 2005; Tykocinski, 2001). According to this idea, disappointing outcomes can prompt people to reason counterfactually about how events could have ended more favorably. Such counterfactual reasoning, in turn, makes people feel worse about the outcome. To avoid feeling worse, people reason that the outcome was “bound to happen,” thereby showing hindsight bias for disappointing self-relevant events (Tykocinski, Pick, & Kedmi, 2002). Interestingly, according to a related concept called defensive processing, people downplay the foreseeability of negative, self-relevant outcomes, thereby resulting in little to no hindsight bias (Louie, 1999; Mark & Mellor, 1991). A literature review in this context yields contradictory findings but also a possible reconciliation.

In two experiments, Tykocinski (2001) found that participants showed more hindsight bias for disappointing outcomes (specifically, negative outcomes for self-relevant events) than for positive ones. In the first experiment, participants who experienced a disappointing outcome (failing to arrive on time for a limited-time discount on a desirable watch) exhibited greater hindsight bias than those who experienced a positive outcome (arriving on time for the discount); that is, those in the former group gave estimates making the disappointing outcome seem inevitable, while those in the latter group gave estimates closer to chance. The second experiment used a real situation instead of an imaginary scenario—the 1999 election for the prime minister of Israel. Consistent with the previous experiment, Tykocinski found that participants who favored the candidate who lost overestimated the probability of the other candidate’s victory—that is, they exhibited greater hindsight bias than participants who favored the candidate who won. These results relate to the impressions of inevitability component (“It had to happen”) from the three-component model of hindsight bias (see Table 1). Participants managed their disappointment by making negative outcomes seem more likely and positive.
Looking Backward and Forward on Hindsight Bias, Bernstein et al. (in press)

outcomes seem less likely; Tykocinski termed this effect “retroactive pessimism.” In a similar study, Tykocinski and Steinberg (2005) compared events in which the negative outcome was a “near miss” (e.g., missing a flight by 10 minutes) as opposed to a “far miss” (e.g., missing a flight by 50 minutes), while varying the consequences of the loss (a $30 vs. a $435 loss). They found that when the stakes were higher, a “near miss” resulted in more hindsight bias than a “far miss.” Hindsight bias here was measured by probability judgments on a series of self-generated counterfactual statements, such as “If I had used the pre-flight service and checked in my suitcase the day before, this would not have happened.” In the context of these experiments, this means that participants also experienced more retroactive pessimism in the “near miss” condition, relative to the “far miss” condition—for example a participant who experienced a near miss in the high stakes condition would feel that early planning (such as checking in their luggage in advance) would have been relatively unlikely to change the outcome, as compared to one who experienced a far miss. The authors argue that using retroactive pessimism to manage disappointment requires suppressing counterfactual thinking to avoid concluding that a positive outcome was likely.

The findings of Mark and Mellor (1991) are an interesting counterpoint to those of Tykocinski; across multiple experiments, Mark and Mellor found that participants showed less hindsight bias for negative outcomes of self-relevant events. Mark and Mellor examined retrospections about the foreseeability of a job lay-off from three groups: laid-off workers, workers unimpacted by the lay-offs, and community members. The authors found that laid-off workers—for whom the event had a negative outcome and was self-relevant—deemed the lay-offs to be less foreseeable and more unexpected than the two other groups. Thus, the laid-off workers experienced the least hindsight bias. Mark, Boburka, Eyssell, Cohen, and Mellor (2003) found additional support for this. They gave two participants hypothetical stock in a stock market decision-making game, while a third participant observed the game without participating. One participant experienced an extreme result—either a major loss or a major gain—while the other participant experienced only a modest gain. The authors found that hindsight bias (in this case, the degree to which the participant felt that the outcome was foreseeable) was lowest for those participants who experienced the major loss, in comparison to both the unaffected participants and to other groups where the participants experienced a major gain; this is the opposite of what retroactive pessimism would predict, where one would expect to see greater hindsight bias for the participants with the most negative outcomes (i.e., the major loss). These examples are consistent with defensive processing (Pezzo & Pezzo, 2007); the idea here is that perceiving the event as less predictable serves to protect the participant from the implication that the outcome was a consequence of their own bad decisions.

Further support for defensive processing comes from a study in which business students playing a stock market game perceived negative outcomes resulting from their decisions as less foreseeable than positive outcomes (Louie, 1999). Similarly, Hörlzl, Kirchler, and Rodler (2002) found that participants who tried to recall their predictions about the success of the euro a year after their original predictions perceived their correct predictions as more foreseeable than their
incorrect predictions. For example, euro supporters—those who supported the introduction of the euro before and after the euro’s introduction—showed more hindsight bias for positive than for negative economic developments associated with the euro’s introduction.

The findings from defensive processing almost perfectly contradict those predicted by retroactive pessimism; Pezzo and Pezzo (2007) combined these findings in a single, parsimonious cognitive theory of hindsight bias, under the umbrella of a sense-making theory that they called “motivated sense-making.” Pezzo and Pezzo (2007) extended Pezzo’s (2003) sense-making model with an eye for self-relevant outcomes, neatly tying into retroactive pessimism (Tykocinski & Steinberg, 2005; Tykocinski, 2001) and defensive processing (Louie, 1999; Mark & Mellor, 1991). Recall that Pezzo (2003) distinguished between initial surprise (where the answer or outcome is incongruent with prior expectations) and resultant surprise (when participants are unable to make sense of a provided answer); Pezzo and Pezzo argue that, for items with high initial surprise and negative self-relevant outcomes, the outcome depends on the ability of the participant to make sense of the correct answer in terms of either external or internal causes.

- In the case that the participant is able to generate an external source for the incongruity with relative ease (i.e., the outcome has high initial surprise but low resultant surprise), he or she experiences strong hindsight bias (Pezzo and Pezzo tie this into retroactive pessimism).

- If this fails (i.e. the outcome has high resultant surprise), the participant experiences little or no hindsight bias, attributing the incongruity to chance or unpredictable factors (which ties into defensive processing).

- However, if internal reasons (i.e., those which must be attributed to the participant) are too obvious to be ignored, then the participant can “accept responsibility” for the incongruity and adjust the perceived likelihood of the outcome accordingly. This results in relatively strong hindsight bias.

- In the preceding case, it is also possible that the participant would choose not to report this hindsight bias publicly, in order to maintain a good image; as previous research has indicated, self-presentation concerns can moderate hindsight bias effects (Campbell & Tesser, 1983; Hawkins & Hastie, 1990; Musch & Wagner, 2007). Note that this case is largely restricted to hindsight bias for outcomes of events, rather than answers to questions; in the latter case, there is no way for the participant to “hide” his or her hindsight bias.

In sum, hindsight bias is a highly robust phenomenon across methodologies and theoretical frameworks. We discussed explanations based on memory, reconstruction, and motivational processes. It is highly likely that all these factors work in concert to produce hindsight bias. Future studies might consider additional potential explanations, as well as specific combinations of factors that produce hindsight bias. The next section provides some future directions that we see as worth pursuing.
Looking Forward

The evidence that scientists have accumulated thus far on hindsight bias highlights many issues to explore. Here, we suggest three areas of future research that we find particularly exciting. The first involves the link between hindsight bias and metacognition. The second involves the link between hindsight bias and theory of mind (see also Löffler & Schneider, this volume), and the third involves the link between hindsight bias and learning.

Hindsight Bias and Metacognition

Although metacognitive explanations of hindsight bias exist, methodologies that are common in the metacognitive literature are rarely used to study hindsight bias. Confidence ratings have been used in the literature, although not extensively (e.g., Hoch & Loewenstein, 1989). This led to the surprising hypercorrection finding that when people have high confidence in their initial wrong response, they can correct their error more successfully than when they have low confidence in their initial wrong responses (Butterfield & Metcalfe, 2001). In such cases, there is a solid basis for the experience that they knew the correct response all along (Metcalfe & Finn, 2011). See Figure 3.

Other associations between confidence and accuracy can expose additional interesting aspects of the answering process in the presence of hindsight bias. In particular, an association taking a central role in the metacognitive literature is resolution. Resolution reflects the extent to which confidence ratings discriminate between correct and incorrect answers and is measured by a within-participant correlation between confidence and accuracy. Reliable discrimination is important for regulatory decisions, such as whether to provide an answer or respond by “I don’t know” (see Werth et al., 2002). This is relevant for contexts such as eyewitness testimony, when one can provide some pieces of information and withhold others, or in educational tests that allow choice among questions. In the context of text learning, individual differences—such as cognitive ability—and conditional factors—such as in-depth processing—correlate with better resolution (Griffin, Wiley, & Thiede, 2008). An example of individual differences is the finding that older children showed better resolution than younger children when answering general knowledge questions (Koriat & Ackerman, 2010a). An example of conditional factors that affect resolution is the finding by Mitchum and Kelley (2010) in the context of problem solving. In their study, participants who attempted to solve a problem by themselves, before seeing the response options (either spontaneously or because of instructions), showed better resolution than those who answered the same problems in a multiple-choice test format, which involves choosing the best among given options. This finding suggests that more in-depth processing occurs when participants generate the solution themselves. Currently, it is unknown whether there is a link between resolution and hindsight bias and how individual differences and conditional factors affect resolution in hindsight bias contexts.
Hindsight Bias and Theory of Mind

Earlier in this chapter, we described the classic hindsight bias task in which people predict the outcome to an event, such as an election or war. Upon learning the outcome, they try to recall their prediction. Most people fail to ignore the outcome and say they “knew it all along” (Fischhoff, 1975; Wood, 1978). This form of hindsight bias found among adults resembles a well-established error in developmental psychology involving children’s failure to acknowledge false beliefs. False belief understanding is an important aspect of what is called theory of mind. In a classic, now-standard task, an experimenter shows a preschooler a familiar container such as a crayon box, and asks the child what is inside. The child reasonably answers, “crayons.” Next, the experimenter opens the box to reveal something unexpected, such as a small toy pig. The experimenter then closes the box, and asks the child what she first thought was inside the box, what is really inside the box, and what another child who hasn’t peeked inside the box will think is inside. Remarkably, most three-year-olds claim that they originally thought that there was a pig inside and that a same-age peer will also think that there is a pig inside. Five-year olds, conversely, answer “crayons” correctly. Importantly, three-year-olds and five-year-olds alike report that there is really a pig inside (Gopnik & Astington, 1988). Thus, three-year-olds aren’t simply confused about the game. Rather they seem to have trouble understanding that beliefs can be false and that other minds differ from their own mind (see Wellman, Cross, & Watson, 2001). Most hindsight bias and theory of mind studies involve adults and children, respectively, although findings demonstrate that assessing others’ knowledge is hard even for adults (e.g., Apperly, Samson, & Humphreys, 2009; Carruthers, 2009; Keysar, Lin, & Barr, 2003; Koriat & Ackerman, 2010b; Koriat, Nussinson, & Ackerman, 2014; Sommerville, Bernstein, & Meltzoff, 2013). Besides surface-level similarities, the two constructs relate in that adults and children fail to set aside what they know when estimating what another person knows (Bernstein, Atance, Loftus, & Meltzoff, 2013; Mitchell et al., 1996; Royzman et al., 2003; M. Taylor, 1988).

There is relatively little work on the link between hindsight bias and theory of mind, in comparison to the huge, separate literatures on each of these constructs. To our knowledge, M. Taylor (1988) was the first to propose a link between hindsight bias and theory of mind, and Mitchell and L. Taylor (1999) provided direct empirical evidence for this link. Mitchell and L. Taylor asked four- to seven-year-old children to look at a circular disk oriented at a slant inside a dark container. When children knew that the disk was really a circle, they judged the disk as more circular than the shape that they could actually see. This bias correlated positively with the number of errors children made on standard theory-of-mind tasks, even after controlling for age differences in theory of mind. Bernstein, Atance, Meltzoff, and Loftus (2007) provided further evidence for this link in preschoolers: three- to five-year old children completed a battery of hindsight bias and theory of mind tasks. Once again, hindsight bias correlated positively with errors on theory of mind tasks, after controlling for various factors known to correlate with theory of mind such as age, language ability, and executive function. Extending this link to school-age children, Massaro and colleagues found mixed evidence for a link between hindsight
bias and more advanced forms of theory of mind (Massaro, Castelli, Sanvito, & Marchetti, 2014; see also Lagattuta, Sayfan, & Harvey, 2013). In these more advanced forms of theory of mind, participants judge the beliefs of two or more characters simultaneously. For example, participants might have to judge what one character in a story thinks that another character in the story thinks. Although it is unlikely that hindsight bias and theory of mind are identical, the two are indeed linked. Future work should aim to explain this link.

Hindsight Bias and Learning

On any given day, parents ask their children what they learned in school. Children often reply, “nothing.” There are many reasons for this answer, but an important one may be that children (M. Taylor, Esbensen, & Bennett, 1994), and even adults, do not remember learning—thinking instead that they “knew it all along.” If children and adults recode new knowledge as “I knew it” (Fischhoff, 1977; Slovic & Fischhoff, 1977), an illusion of understanding may lead to a failure to learn from the past (Fischhoff, 1982; Henriksen & Kaplan, 2003; Roese & Vohs, 2012). Others view hindsight bias as adaptive (Hawkins & Hastie, 1990; Hoffrage et al., 2000; see also Schacter, Guerin, & St. Jacques, 2011; Sommerville & Hammond, 2007) and argue that it is better to update knowledge with the correct answer than confuse one’s original answer with the correct answer.

These two views of hindsight bias, that it results in failure to remember learning and that it is a form of adaptive learning, seem to imply that hindsight bias should aid learning. While the failure-to-remember view addresses whether we keep track of changing knowledge, the adaptive-learning view addresses how we acquire new knowledge. Returning to the children who think that they learn nothing in school, are these children better or worse learners than children who think that they do learn in school? Similarly, does thinking that one already knows everything make someone less open to learning new things? Perhaps know-it-alls or “knowists” as the first author’s niece calls such people, use less effective study strategies than people who realize that they don’t know it all. Despite nearly 40 years of research, little work has examined hindsight bias’s effects on learning (Blank et al., 2007; see Appleton-Knapp & Bjork, 2006; Metcalfe & Finn, 2011). Ongoing work of ours shows that hindsight bias may relate to learning, but that it depends on how one measures hindsight bias. For example, when we measured hindsight bias on a continuous scale using open-ended almanac questions requiring a numerical answer (e.g., “How many keys are there on a standard piano?”), we found no correlation between hindsight bias and learning (Bernstein, Aujla, Erdfelder, & Peria, 2011). Conversely, when we measured hindsight bias on a categorical scale using multiple-choice answers (e.g., In the Stroop Effect, what task is automatized? A: naming colors; B: naming shapes; C: reading; D: all of the above), we found a positive correlation between hindsight bias and learning (Le Grand, Bernstein, Kumar, & Butler, 2013). Note that in both the continuous and categorical measures of hindsight bias, we operationalized learning as the ability to recall the correct answer when asked for the correct answer at the end of the experiment. As with the link between hindsight bias and theory of mind, more work is needed to explain the link between hindsight bias and learning.
This latter link also relates more broadly to the emerging field of learning science (see Dell’Antonia, 2014; Roediger & Karpicke, 2006).

Final Word

In this chapter, we demonstrated that hindsight bias is a complex cognitive error. As explained in the introduction, this phenomenon has practical implications in real life contexts, beyond the psychological laboratory, such as in law and education. Understanding the underlying processes can guide future attempts to attenuate hindsight bias and delineate the conditions under which hindsight bias is absent. As with most topics in psychology, we call for future studies to continue these intriguing lines of research and contribute to solving at least some of the puzzles.

Acknowledgment

We thank the Canada Research Chairs Program (950-228407) for funding this work.

References


tree model for separating recollection and reconstruction in hindsight. Journal of Experimental

Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under
uncertainty. Journal of Experimental Psychology: Human Perception and Performance, 1, 288–
299.

Human Perception and Performance, 3, 349–358.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under
Press.


Fischhoff, B., & Beyth, R. (1975). “I knew it would happen”: Remembered probabilities of once-
future things. Organizational Behavior and Human Performance, 13, 1–16.


its relation to the understanding of false belief and the appearance-reality distinction. Child
Development, 59, 26–37.

explanation: Concurrent processing and cue validity as constraints on metacomprehension

Groß, J., & Bayen, U. (2015). Hindsight bias in younger and older adults: The role of access


Harley, E. M., Carlsen, K., & Loftus, G. R. (2004). The “saw it all along” effect: Demonstrations of
visual hindsight bias. Journal of Experimental Psychology: Learning, Memory, & Cognition, 30,
960–968.

Learning and Verbal Behavior, 20, 86–96.

outcomes are known. Psychological Bulletin, 107, 311–327.


Looking Backward and Forward on Hindsight Bias, Bernstein et al. (in press)


The Telegraph (2012).


